8E

УДК 004.93

А.А. Егоров
Днепропетровский национальный университет им. О. Гончара,
г. Днепропетровск, Украина
egorov@mayak.dp.ua

Повышение быстродействия метода гибридной нечеткой кластеризации за счет динамического сжатия размерности карты Кохонена

В статье предложен модифицированный алгоритм гибридной нечеткой кластеризации sFCM, в котором применяется динамическое сжатие размерности карты Кохонена, что позволяет снизить время обучения. Представлены экспериментальные результаты применения предложенного алгоритма для кластеризации низкоконтрастных полутоновых медицинских изображений.

Введение

В настоящее время при обработке данных приходится часто сталкиваться с необходимостью выполнения кластеризации, для осуществления которой разработано множество различных методов. Однако решение этой задачи часто осложняется такими факторами, как большой объем исходных данных и отсутствие априорных сведений о скрытых в них зависимостях. Первый фактор может накладывать на используемые методы кластеризации ограничения, связанные с быстродействием, а негативное влияние второго фактора обусловлено неизвестным числом кластеров, что является важнейшим управляющим параметром у большинства алгоритмов и неоднозначностью разделения исходных данных.

Таким образом, интерес представляют методы кластеризации, позволяющие:

– осуществлять подстройку числа кластеров в процессе работы;

– учитывать неоднозначность разделения исходных данных между кластерами, что присуще, например, нечетким системам;

– использовать самоорганизующиеся алгоритмы.

Одним из способов подстройки числа кластеров в процессе работы является задание их избыточного числа с дальнейшим уменьшением их количества при обучении. Однако этот подход приводит к существенному возрастанию времени кластеризации, что неприемлемо, особенно при обработке данных в реальном режиме.

Следовательно, интерес представляют такие алгоритмы кластеризации, которые позволяют, с одной стороны, обходиться без жесткого задания числа кластеров, учитывать неоднозначность разбиения данных и обучаться без учителя, а с другой стороны, – иметь компенсационные механизмы, направленные на снижение негативного влияния от возрастания вычислительных затрат в случае введения избыточного числа кластеров.

В настоящее время существуют методы кластеризации, которые удовлетворяют ряду из приведенных выше требований. Примером таких методов являются нечеткие алгоритмы, позволяющие учитывать неоднозначность разбиения исходных

данных между кластерами. Одним из активно применяемых методов нечеткой кластеризации является алгоритм FCM (Fuzzy c-means), что связано как с его простотой, так и с достаточным для большинства задач уровнем чувствительности. При выполнении нечеткой кластеризации вводятся понятия нечетких кластеров и функции принадлежности объектов к ним, изменяющейся на отрезке [0,1], что позволяет оценить степень принадлежности объекта к тому или иному кластеру [1]. Однако недостатком алгоритма FCM является необходимость задания неизменного числа нечетких кластеров.

Другим часто используемым в настоящее время подходом к решению задачи кластеризации является применение нейронных сетей, обучаемых без учителя. Примером такой сети может служить карта Кохонена. При ее использовании задается двумерная сетка из нейронов, веса которых меняются в процессе обучения. Изменение происходит как у весов нейрона-победителя (определяется на основании минимума Евклидова расстояния до векторов исходных данных), так и у соседних нейронов [2]. Однако недостатками этого подхода является отсутствие учета неоднозначного разделения данных между нейронами в явном виде и необходимость задания размерности сети перед началом обучения.

Постановка задачи

Перспективным подходом являются нейро-фаззи алгоритмы, представляющие собой объединение нечеткой кластеризации и нейронной сети в пределах одного метода. Примером реализации такого подхода может служить алгоритм гибридной нечеткой кластеризации sFCM, представленный в работе [3]. Применение этого метода позволяет повысить чувствительность кластеризации и уменьшить время ее выполнения. Однако базовый алгоритм sFCM также требует задания начального значения неизменного числа нечетких кластеров. В работе [4] была предложена модификация алгоритма sFCM, в которой применяется динамическое сжатие функции принадлежности, что приводит к уменьшению их числа. Однако такой подход требует, чтобы начальное число нечетких кластеров было задано с избытком, позволяющим выполнить их последующее динамическое уменьшение. Это приводит к повышению вычислительной нагрузки, что особенно заметно при увеличении числа нечетких кластеров перед применением сети Кохонена.

Целью данной статьи является снижение временных затрат на работу метода гибридной нечеткой кластеризации sFCM, что достигается благодаря его модификации за счет двухэтапного динамического сжатия размерности карты Кохонена.

Решение задачи

Предложенный модифицированный метод гибридной нечеткой кластеризации sFCM состоит из следующих шагов:

1. Инициализация числа кластеров c, значения m (экспоненциальный вес нечеткой кластеризации), начальных значений центров нечетких классов v_{fcm}^{0} (например, путем равномерного заполнения значениями векторов исходных данных).

2. Формирование начальных значений весов нейронов карты Кохонена, размерность которой выбирается равной $[N*c,1]$, где N – коэффициент увеличения количества кластеров (рекомендуемые значения – 2 или 4). Каждый новый центр нечетких кластеров получается на основе значений из матрицы v_{fcm}^{t-1} центров предыдущей итерации методом пропорционального распределения [5].

Вычисляется матрица Евклидовых расстояний в q-мерном пространстве между нечеткими кластерами.

Производится упорядочение центров кластеров, в результате чего формируется вектор d , содержащий расстояния между соседними центрами.

Между каждой парой центров нечетких кластеров формируются новые центроиды по следующему алгоритму:

а) вычисляется вектор Δ:

$$\Delta_j = \left(v_{fcm}^{t-1}\right)_{k,j} - \left(v_{fcm}^{t-1}\right)_{k+1,j}, k \in \{1,...,c-1\}, j \in \{1,...,q\}, \quad (1)$$

где q – количество информативных признаков, описывающих каждый экземпляр данных;

б) значение N_k^{ctr} уменьшается на 1, пока выполняются условия:

$$N_k^{ctr} > 0 \text{ и } \left|\Delta_j / \left(N_k^{ctr} + 1\right)\right| \le \Delta d_{\min}, \quad (2)$$

причем каждый элемент вектора N^{ctr} содержит число новых центров кластеров между парами исходных центроидов. Эти значения прямо пропорциональны расстоянию между соседними исходными центрами:

$$N_k^{ctr} = (N-1) \times c \times \frac{d_{k,k+1}}{\Delta d_{sum}} \times \left(1 + \frac{d_{c-1,c}}{\Delta d_{sum}}\right), \quad (3)$$

где Δd_{sum} – сумма элементов вектора d , а $\Delta d_{\min}$ вычисляется так:

$$\Delta d_{\min} = d_{\min} / \left(2 \times N \times c \times N_{\max}^{ctr}\right), \quad (4)$$

причем $d_{\min}$ – минимальный элемент вектора d , а $N_{\max}^{ctr}$ – максимальный элемент вектора N^{ctr} ;

в) если $N_k^{ctr} > 0$, то формируется очередной новый вектор центров $\left(v_{som}^t\right)^0$:

$$\left(v_{som}^t\right)_{L,j}^0 = \left(v_{fcm}^{t-1}\right)_{k,j} + \Delta_j / \left(N_{k,k+1}^{ctr} + 1\right), L \in \{1,...,N \times c\}. \quad (5)$$

В случае, когда Евклидово расстояние от L -го центра до одного из предшествующих ему $L-1$ центроидов не превышает $\Delta d_{\min}$, этот центр удаляется из вектора $\left(v_{som}^t\right)^0$. Это замечание касается и исходных центров кластеров. Таким образом происходит первоначальное уменьшение количества нейронов, количество которых не может быть меньше, чем c .

3. Кластеризация исходных данных картой Кохонена, в процессе которой происходит динамическое сжатие размерности сети:

3.1 Инициализация управляющих параметров нейронной сети. При этом в качестве начальных значений весов нейронов используется векторы $\left(v_{som}^t\right)^0$. На основании многочисленных экспериментальных исследований рекомендуется использовать следующие управляющие параметры:

– в качестве функции соседства выбирать функцию Гаусса;

– нейроны располагать в столбец;

– формировать вектор радиусов влияния каждого нейрона из натуральных

чисел от r_1^0 до 1, причем:

$$r_1^0 = N_n^0 / 3, \qquad (6)$$

где N_n^0 – начальное количество нейронов;

– количество шагов обучения $i_{\max}$ выбирать равным длине вектора радиусов.

3.2 Вычисление значения $p_{\max}$ по следующим формулам:

$$p_{\max} = \left(p_{\max}^0\right)^{2-0.5/p_{\max}^0}, \qquad (7)$$

$$p_{\max}^0 = \left(N_n^0\right)_{\min} / N_n^0, \qquad (8)$$

причем $\left(N_n^0\right)_{\min}$ – количество нейронов после предварительного сжатия, выполняемого согласно пунктам 6.1 – 6.4, 6.7, 6.8. При этом $N_c = c/\left(2\times N_n^0\right)$, а минимально допустимое число нейронов $c_{\min} = c$.

3.3 Каждый шаг обучения карты Кохонена содержит следующие действия:

3.3.1 попытка уменьшения числа нейронов, если:

$$i / i_{\max} < p_{\max}, \qquad (9)$$

где i – номер текущей итерации обучения, причем $i \le i_{\max}$. При выполнении сжатия $N_c = \dfrac{c\times p_{\max}}{2\times N_n^0}$, а $c_{\min} = c + \left(\dfrac{N_n^0 - c}{i_{\max}}\right)\times\left(i_{\max} - i\right)$. В случае уменьшения количества нейронов вычисляется значение r_1^i по формуле (6) и разность $\Delta r_1 = r_1^{i-1} - r_1^i$. Если $\Delta r_1 > 0$, то номер текущей итерации увеличивается на значение Δr_1;

3.3.2 на основе Евклидовых расстояний для каждого экземпляра исходных данных определяется ближайший к нему нейрон (нейрон-победитель);

3.3.3 происходит модификация весов нейронов-победителей и соседних с ними нейронов, определяемых с помощью выбранной функции соседства;

3.3.4 переход к следующей итерации, если $i < i_{\max}$.

4. Получение новых значений центров нечетких кластеров v_{som}^t путем выбора c значимых центров из матрицы весов нейронов, полученной в результате обучения карты Кохонена. Этот выбор является нетривиальной задачей и оказывает влияние на чувствительность метода.

5. Вычисление текущих значений функции принадлежности u:

$$u_{k,i} = \left[\sum_{L=1}^{c}\left[\frac{\left(\sum_{j=1}^{q}\left(X_{i,j} - (v_{som}^t)_{k,j}\right)^2\right)^{1/2}}{\left(\sum_{j=1}^{q}\left(X_{i,j} - (v_{som}^t)_{k,L}\right)^2\right)^{1/2}}\right]^{\frac{2}{m-1}}\right]^{-1} \left(\begin{matrix}\forall k \in \{1,\ldots,c\},\\ \forall i \in \{1,\ldots,n\}\end{matrix}\right), \qquad (10)$$

где n – число экземпляров данных, а X – исходные данные.

6. Динамическое сжатие функции принадлежности на основе расстояний между центрами нечетких кластеров происходит итеративно по следующему алгоритму:

6.1 определение матрицы расстояний в q-мерном пространстве между центрами нечетких кластеров d^0 и значений минимального ненулевого $d^0_{\min}$ и максимального $d^0_{\max}$ расстояний;

6.2 автоматическое определение значений d^1_a и d^2_a – минимально допустимых расстояний между центрами кластеров – на основе статистических характеристик и по изменению производной, соответственно.

$$d^1_a = d^1_c + \left(\sum_{i=1}^{n^1_d} \left(d^1_c - d^1_i \right) \right) / n^1_d , \quad (11)$$

$$d^1_c = \min\left(\overline{d^1}, 0.5 \times \max(d^1) \right), \quad (12)$$

где $\overline{d^1}$ – среднее по вектору d^1, который состоит из n^1_d элементов d^0_i матрицы d^0, удовлетворяющих условию:

$$d^0_i < d^0_c + \left(\sum_{j=1}^{n^0_d} \left(d^0_c - d^0_j \right) \right) / n^0_d , \quad (13)$$

$$d^0_c = \min(\overline{d^0}, 0.5 \times d^0_{\max}), \quad (14)$$

где $\overline{d^0}$ – среднее по матрице d^0. Расстояние d^2_a вычисляется следующим образом:

а) на основании матрицы d^0 формируется отсортированная по возрастанию последовательность расстояний;

б) для каждого элемента в последовательности вычисляется его производная по отношению к 1-у элементу и производится поиск минимального значения этой производной;

в) значение d^2_a вычисляется как среднее между элементом последовательности с минимальным значением производной и следующим за ним элементом;

6.3 выбор минимального ($d_{\min}$) из расстояний d^1_a, d^2_a и $\left(d^0_{\max} - d^0_{\min} \right) \times N_c$, что позволяет избежать негативного влияния от использования слишком большого значения параметра N_c, который может изменяться от 0 до 1, причем при значениях N_c, больших 0.2, как правило, автоматически выбирается d^1_a или d^2_a. Если $d_{\min} = 0$ или не существует расстояния между двумя нечеткими кластерами, меньшего $d_{\min}$, то переход к пункту 6.9;

6.4 изменение каждого центра кластера, отличного от двух выбранных:

$$\left(v^t_{fcm} \right)_k = \left(v^t_{fcm} \right)_k \times (1 - \Delta_d) + \left(v^t_{fcm} \right)_d \times (\Delta_d), \quad (15)$$

$$\Delta_d = \left(d^0 \right)_{d,s} / \left(d^0 \right)_{d,k}, (k \neq d, k \neq s), \quad (16)$$

где d, s – индексы выбранных для сжатия удаляемого и модифицируемого центров, соответственно; $\left(d^0\right)_{d,s}$ – расстояние между выбранными центрами; $\left(d^0\right)_{d,j}$ – расстояние между удаляемым и одним из оставшихся центров. По формуле (15) модифицируются и центры v_{fcm}^{t-1};

6.5 изменение функции принадлежности $u_{k,i}$ к каждому из кластеров для каждого из векторов исходных данных:

$$u_{k,i} = \max\left(u_{k,i} \times (1-\Delta_d), u_{d,i} \times (\Delta_d)\right), \tag{17}$$

где $u_{d,i}$ – функция принадлежности к удаляемому кластеру;

6.6 изменение функции принадлежности (путем выполнения нечеткой операции объединения) для двух выбранных центров кластеров, которое отображается для модифицируемого центра. Затем происходит сжатие путем отбрасывания функций принадлежности удаляемого кластера;

6.7 преобразованию, описанному в пункте 6.6, подвергаются векторы центров нечетких кластеров v_{som}^t и v_{fcm}^{t-1}, причем изменение модифицируемого центра кластера происходит с помощью усреднения;

6.8 если количество нечетких кластеров, полученных после сжатия, больше, чем $c_{\min}$ (как правило, равно 2), то переход к пункту 6.1, а иначе – к пункту 6.9;

6.9 масштабирование полученной после сжатия функции принадлежности к нечетким кластерам для каждого экземпляра исходных данных, чтобы удовлетворялось условие $\sum_{k=1}^{c} u_{k,i} = 1$.

7. Вычисление значений векторов центров нечетких кластеров v_{fcm}^t, которые будут использованы в начале следующей итерации:

$$(v_{fcm}^t)_{k,j} = \left(\sum_{i=1}^{n} \left(u_{k,i}\right)^m \cdot X_{i,j}\right) / \sum_{i=1}^{n} \left(u_{k,i}\right)^m . \tag{18}$$

8. Если выполняются условия:

$$\frac{\sum_{k=1}^{c} \sqrt{\sum_{j=1}^{q} \left(\left(v_{fcm}^t\right)_{k,j} - \left(v_{fcm}^{t-1}\right)_{k,j}\right)^2}}{c} > \varepsilon , \tag{19}$$

$$\left|V_{xb}^t - V_{xb}^{t-1}\right| > \varepsilon \text{ или } \left|V_{fz}^t - V_{fz}^{t-1}\right| > \varepsilon , \tag{20}$$

где ε – пороговое значение, а V_{xb}^t, V_{xb}^{t-1} и V_{fz}^t, V_{fz}^{t-1} – показатели Ксие-Биени и нечеткости текущей и предыдущей итераций соответственно, которые вычисляются по следующим формулам [6]:

$$V_{xb} = \left(\sum_{k=1}^{c} \sum_{i=1}^{n} \left(u_{k,i}\right)^m \times \sum_{j=1}^{q} \left(X_{i,j} - (v_{fcm}^t)_{k,j}\right)^2\right) / \left(n \times \left(d_{\min}\right)^2\right), \tag{21}$$

$$V_{fz} = \left(\sum_{k=1}^{c} \sum_{i=1}^{n} \left(u_{k,i} \right)^m \right) / n, \qquad (22)$$

причем $d_{\min}$ – минимальное расстояние между центрами нечетких кластеров, то выполняется переход к пункту 2.

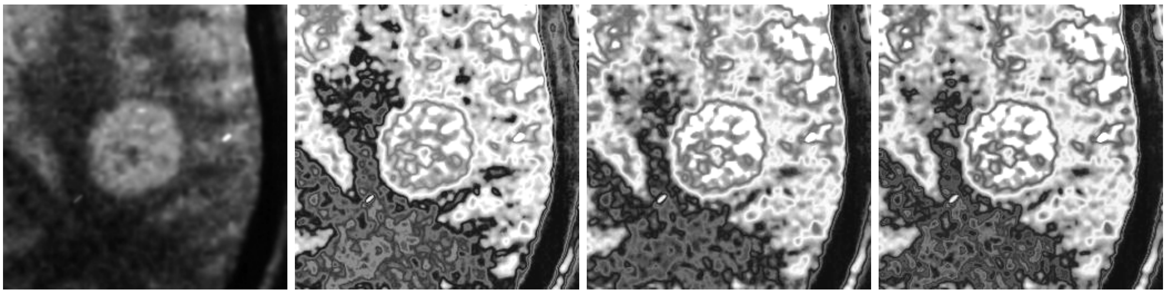
Таким образом, в предложенном методе кластеризации происходит двухэтапное сжатие размерности нейронной сети: предварительное – перед началом кластеризации картой Кохонена при формировании новых центров нечетких кластеров для повышения чувствительности, и динамическое – в процессе обучения.

Экспериментальные результаты были получены при обработке различных низкоконтрастных изображений, в том числе медицинских, примером которых служат полутоновые снимки, представленные на рис. 1а (спин-решетчатая Т1 релаксация ЯМР (ядерного магнитного резонанса) участка головного мозга) и рис. 2а (результаты томограммы).

При кластеризации приведенных на рис. 1 а и 2 а изображений использовались следующие значения управляющих параметров: $c = 20$ (выбиралось с запасом, чтобы благодаря динамическому сжатию функции принадлежности определить рекомендуемое количество нечетких кластеров); $N = 4$; после применения карты Кохонена выбирались центры кластеров с максимальным количеством относящихся к ним (имеющих минимальное Евклидово расстояние) векторов исходных данных; $N_c = 0{,}5$ (приводит к автоматическому формированию минимально допустимого расстояния между кластерами при сжатии); сжатие выполнялось на основе матрицы Евклидовых расстояний. Визуализация результатов нечеткой кластеризации производилась на основе сравнения с исходными данными [7].

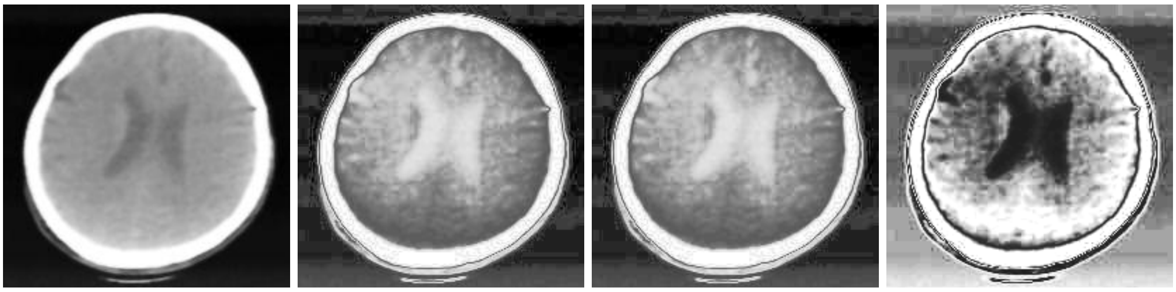
При кластеризации методом sFCM снимка, приведенного на рис. 1а, как без уменьшения числа нейронов (рис. 1б), так и с предварительным снижением размерности нейронной сети (рис. 1в) и двухэтапным сжатием (рис. 1г), не было выявлено существенных различий в результирующих изображениях, что свидетельствует о неизменном уровне чувствительности. Однако для получения изображений, приведенных на рис. 1в и 1г, было затрачено на 18 и 42% меньше времени, соответственно, чем на формирование изображения на рис. 1б. Это достигалось благодаря снижению числа нейронов (в среднем, на 9 и 40% для изображений, представленных на рис. 1 в и 1 г, соответственно), что приводило к пропорциональному уменьшению количества итераций обучения сети Кохонена.

При кластеризации методом sFCM снимка, приведенного на рис. 2а, как без уменьшения количества нейронов (рис. 2б), так и с предварительным снижением размерности нейронной сети (рис. 2в), не было выявлено различий в полученных после обработки изображениях, что свидетельствует о неизменном уровне чувствительности. Однако применение двухэтапного сжатия (рис. 2г) позволило повысить чувствительность за счет более четкого выделения гематомы и области ее влияния. При этом для получения изображений, приведенных на рис. 2 в и 2 г, было затрачено на 23 и 46% меньше времени соответственно, чем на формирование изображения на рис. 2 б. Это достигалось благодаря снижению числа нейронов (в среднем, на 7 и 39% для изображений, представленных на рис. 2в и 2г, соответственно), что приводило к пропорциональному уменьшению количества итераций обучения сети Кохонена.



а б в г

Рисунок 1 – Кластеризация медицинского изображения: а – исходный снимок (256 х 256 пикселей); результаты кластеризации методом sFCM: б – без уменьшения числа нейронов карты Кохонена; в – с первоначальным; г – двухэтапным сжатием



а б в г

Рисунок 2 – Кластеризация медицинского изображения: а – исходный снимок (204 х 201 пикселей); результаты кластеризации методом sFCM: б – без уменьшения числа нейронов карты Кохонена; в – с первоначальным; г – двухэтапным сжатием

Следует заметить, что эффект сжатия размерности карты Кохонена отчетливее заметен на начальных итерациях работы алгоритма sFCM, пока количество нечетких кластеров еще не уменьшено. В рассмотренных примерах сжатие в этих случаях достигает 58 и 64% для одно- и двухэтапного сжатия соответственно, при кластеризации изображения, представленного на рис. 1а, и 30 и 70% соответственно для изображения, приведенного на рис. 2а.

Выводы

Благодаря осуществлению двухэтапного динамического сжатия размерности карты Кохонена в предложенном модифицированном алгоритме гибридной нечеткой кластеризации sFCM достигается существенное снижение временных затрат без потери чувствительности. При этом в некоторых случаях удается добиться повышения не только быстродействия, но и чувствительности за счет удаления несущественных нейронов. Эффект от уменьшения количества нейронов особенно заметен на первых итерациях обучения, пока не выполнено динамическое сжатие нечеткой функции принадлежности.

Литература

1. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH / Леоненков А. – С.-Пб: БХВ–Петербург, 2003. – 719 с.
2. Кохонен Т. Самоорганизующиеся карты / Кохонен Т. ; [пер. 3 англ. изд. В.Н. Агеева ; под ред. Ю.В. Тюменцева]. – М. : Бином. Лаборатория знаний, 2008. – 665 с.

3. Ахметшина Л.Г. Сегментация низко контрастних изображений алгоритмом гибридной кластеризации SOM-FCM / Л.Г. Ахметшина, А.А. Егоров // Системні технології. – Дніпропетровськ, 2008. – Вип. 2 (55). – С. 34-40.
4. Ахметшина Л.Г. Динамическое сжатие функции принадлежности на основании расстояний между центроидами в алгоритме гибридной нечеткой кластеризации. / Л.Г. Ахметшина, А.А. Егоров / Прикладна геометрія та інженерна графіка. – 2010. – Вип. № 84, т. 1. – С. 88-92.
5. Ахметшина Л.Г. Повышение чувствительности гибридной нечеткой кластеризации на основе формирования центроидов пропорционально расстояниям в q-мерном пространстве / Л.Г. Ахметшина, А.А. Егоров // Геометричне та комп'ютерне моделювання. – 2009. – Вип. 24. – С. 193-198.
6. Рутковский Л. Методы и технологии искусственного интеллекта / Рутковский Л. – М., 2010. – 600 с.
7. Егоров А.А. Визуализация результатов нечеткой кластеризации на основе сравнения с исходными данными. / А.А. Егоров // Матеріали VI міжнародної науково-практичної конференції : «Математичне та програмне забезпечення інтелектуальних систем». – 2008. – С. 117-118.

Literatura

1. Leonenkov A. Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH. S.P.: BHV-Peterburg. 2003. 719 s.
2. Kohonen T. Samoorganizujushhiesja karty. M.: Binom. Laboratorija znanij. 2008. 665 s.
3. Ahmetshina L.G. Systemni tehnologii. Dnipropetrovs'k. Vyp 2 (55). 2008. S. 34-40.
4. Ahmetshina L.G. Prykladna geometrija ta inzhenerna grafika. Vyp 84. T 1. 2010. S. 88-92.
5. Ahmetshina L.G. Geometrychne ta komp'juterne modeljuvannja. Vyp 24. 2009. S. 193-198.
6. Rutkovskij L. Metody i tehnologii iskusstvennogo intellekta. M. 2010. 600 s.
7. Egorov A.A. Materiali VI mizhnarodnoi naukovo-praktichnoi konferencii "Matematechne ta programne zabezpechennja intelektual'nih sestem". 2008. S. 117-118.

А.О. Єгоров

Підвищення швидкодії в методі гібридної нечіткої кластеризації за рахунок динамічного стиснення розмірності карти Кохонена

У статті запропонований модифікований алгоритм гібридної нечіткої кластеризації sFCM, в якому застосовується динамічне стиснення розмірності карти Кохонена, що дозволяє знизити затрати часу на навчання. Представлені експериментальні результати застосування запропонованого алгоритму для кластеризації низькоконтрастних напівтонових медичних зображень.

A.A. Yegorov

The Performance Improvement of the Hybrid Fuzzy Clustering Method Based on Dynamical Compressing of Kohonen SOM Dimension

This article deals with the description of the modified hybrid fuzzy clustering algorithm SFCM, which is used Kohonen SOM dimension dynamical compressing, to deliver learning time decreasing. The experimental results of the application of proposed algorithm for low-contrast halftone image clustering are shown.

Статья поступила в редакцию 10.07.2011.